

The graphic features a white background with stylized red and blue curved stripes on the left side, resembling the American flag. A series of grey stars is arranged in a curved path along the top right. At the bottom, a horizontal row of grey stars spans the width of the page.

# **US PRESIDENTIAL ELECTIONS 1976-2020**

By: Zoe Williams

# OVERVIEW

- This dataset contains 4287 columns and 15 rows.
- The variables collected: year, state, state po, state fips, state cen, state ic, office, candidate, party detailed, write in, candidate votes, total votes, version, notes, and party simplified.
- This dataset can be used to analyze the trends and patterns in presidential elections by exploring relationships between variables such as political parties, total votes, year, and location.

# DATA CLEANING + WRANGLING

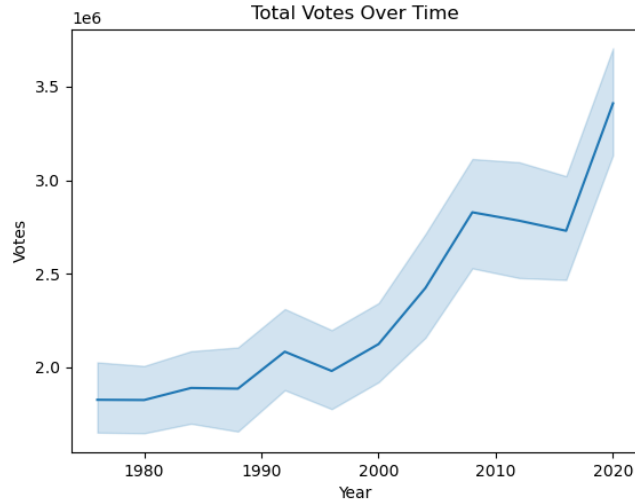
- Four rows contained null values
  - Candidate
  - Party detail
  - Write in
  - Notes
- The fillna function was used to replace the null values with unknown
- The notes row was dropped because all of the values were null
- A vote percentage column was added
  - This was calculated by dividing candidate votes by total votes and multiplying that number by 100

# EXPLORATORY DATA ANALYSIS

- The groupby and sum function was used to determine the total votes cast each year. Result:
  - **1976** 605944064
  - **1980** 663902096
  - **1984** 609936856
  - **1988** 537099170
  - **1992** 770486377
  - **1996** 728343795
  - **2000** 783441739
  - **2004** 768259747
  - **2008** 992684830
  - **2012** 879479158
  - **2016** 941573717
  - **2020** 1865852281
- The value\_counts function was used to determine the most common office contested. Result:
  - **Office US PRESIDENT** 4287
- The groupby and sum function was used to determine the party wise vote distribution. Result:
  - **DEMOCRAT** 1344886700
  - **LIBERTARIAN** 1177252462
  - **OTHER** 6284935371
  - **REPUBLICAN** 1339929297

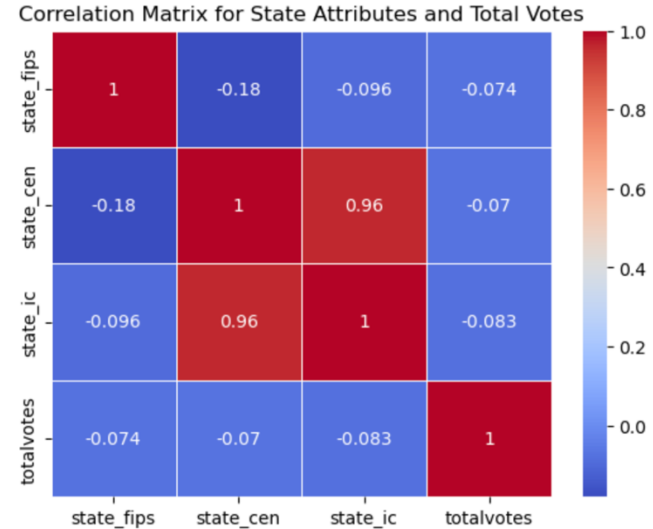
# DATA VISUALIZATION

Line plot: Total Votes Over Time



Although the graph is not a perfect linear line, year and total votes exhibit a positive relationship. There are multiple dips over time, such as 1993-1996 and 2008 to 2012, but there are strong increases from 1995 to 2008 and 2015 to 2020.

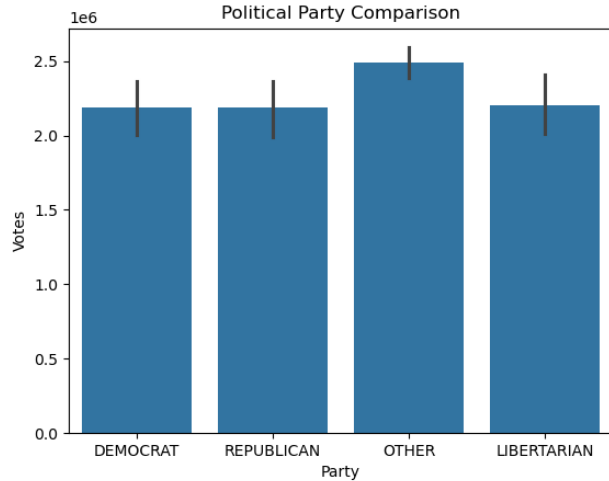
Heatmap: State Attributes and Total Votes



The heatmap demonstrates that there is little to no correlation between state attributes and total votes. On the plot, total votes have negative values and blue boxes for all of that state attribute variables. This could allude to location having no impact on total votes

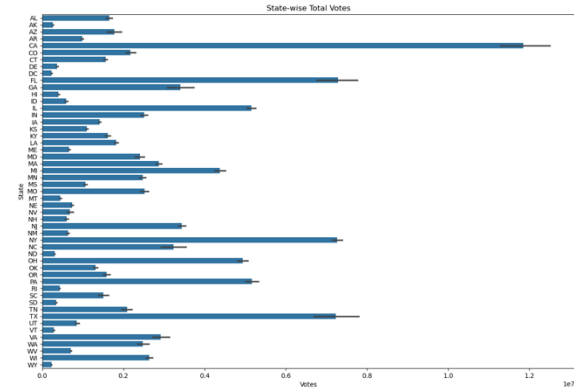
# DATA VISUALIZATION

Bar Graph: Party Comparison



This graph demonstrates that democrat, republican, and libertarian all roughly received the same number of votes. Other, is noticeably higher than the other parties, however this variable contains a large variety of parties. This graph alludes to there being no dominant or more successful party in the two-party system

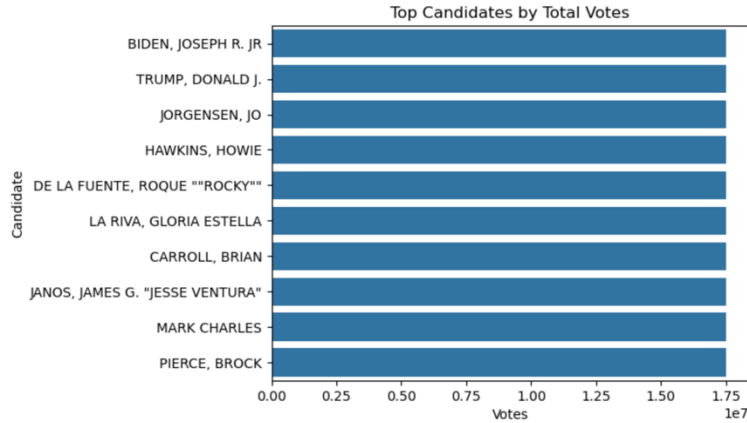
Bar Graph: State-wise total votes



This graph shows that the states with the highest votes are California, Texas, and Florida. The states with the lowest votes are Wyoming, Alaska, and Vermont. These results show that population is strongly correlated with total votes.

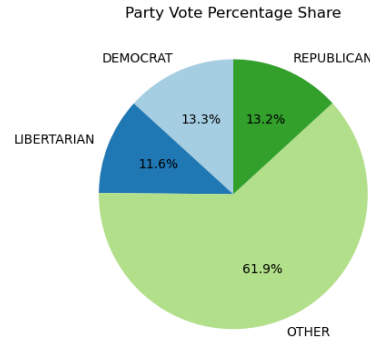
# DATA VISUALIZATION

Bar Graph: Top Candidates



The graph shows that all of the top ten candidates roughly received the same number of votes. This could possibly be the number when 270 electoral votes were met or a mistake in the dataset. This bar graph shows an anomaly in the dataset.

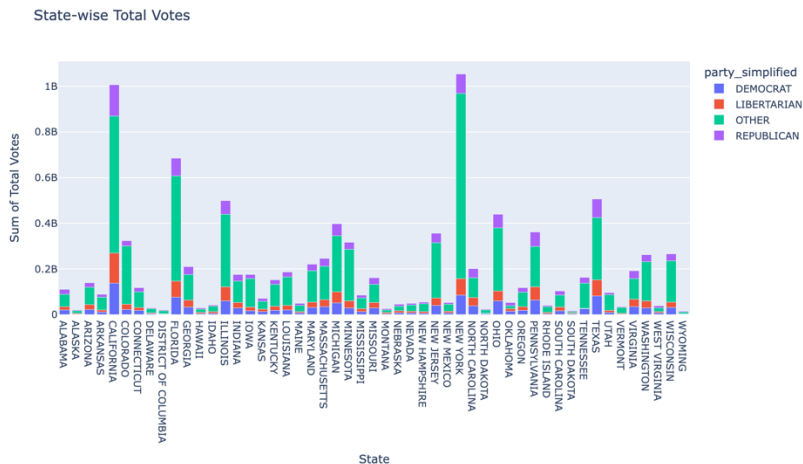
Pie Chart: Party Vote Percentage Share



The pie chart shows that democrats received 13.3%, republicans received 13.2%, libertarians received 11.6%, and other political parties hold 61.9% of the vote. These numbers show that the party's within the two-party system hold nearly the same amount of votes, while smaller political parties dominate a large share of total votes.

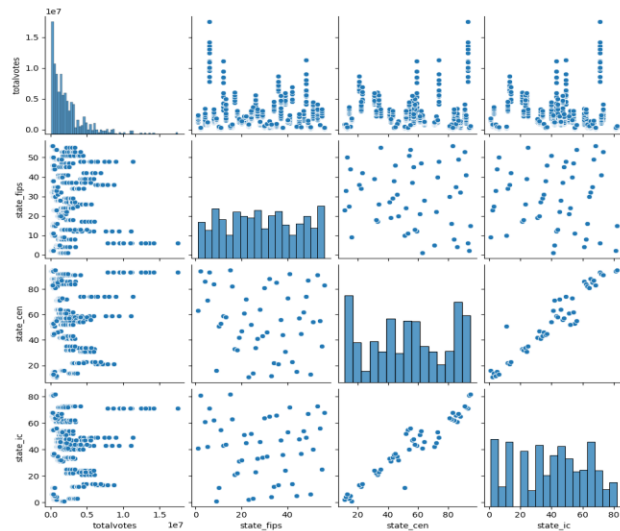
# EXPECTED OUTPUT:

## Histogram: State-wise Total Votes and Political Party



The bar graph clearly shows which states have high votes. The green color is the other party category, and it makes up the majority of many of the state's bars, for most states it appears that republican, democrats and libertarian are around the same size.

## Pairplot: State Attributes and Total Votes

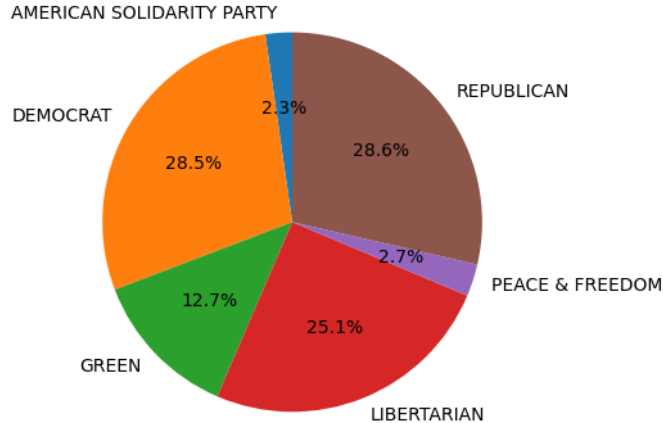


The state attributes and the total votes scatter plots appear clustered near the y-axis or x-axis. This alludes to a weak correlation between the variables.



# ADDITIONAL FINDINGS

Top Political Parties Portion of Total Votes



Due to the overwhelming amount of total vote percentage the other political parties received; my additional graph is a pie chart breaking down some of the parties. . The top three identified parties were selected, along with the top three parties from the other category.

# CONCLUSION

- One of the most prominent findings was the overall growth in total votes, particularly in the years following 1992, with a sharp increase in voter turnout seen in the 2020 election.
- The analysis of state-wise total votes revealed that the most populous states contributed the most to the total vote count
- The major political parties (Democrat, Republican, and Libertarian) received similar levels of support
- The analysis of top candidates by vote share illustrated an anomaly, where many candidates appeared to have received identical vote totals.